ISSN 1324-5910

All correspondance to:

Dr Andrew Worthington
Editor, *Discussion Papers in Economic, Finance and International Competitiveness*
School of Economics and Finance
Queensland University of Technology
GPO Box 2434, BRISBANE QLD 4001, Australia

Telephone: 61 7 3864 2658
Facsimilie: 61 7 3864 1500
Email: a.worthington@qut.edu.au

Queensland University of Technology

DISCUSSION PAPERS IN ECONOMICS, FINANCE AND INTERNATIONAL COMPETITIVENESS

Output-Driven Funding and Budgeting Systems in the Public Sector

Marc Robinson

Discussion Paper No. 101, February 2002

Series edited by
Dr Andrew Worthington

School of Economics and Finance

OUTPUT-DRIVEN FUNDING AND BUDGETING SYSTEMS IN THE PUBLIC SECTOR

Marc Robinson*

Abstract

Output-driven funding systems are systems in which payments made to service-delivery agencies by government are an explicit function of quantities of outputs delivered by those agencies. This paper considers the feasibility of such systems for the funding tax-financed public services. It focuses upon the implications of key characteristics of public sector outputs, and specifically upon the prevalence of heterogeneous outputs, the predominance of services (as opposed to physical goods), and the presence of many 'contingent capacity services'.

Introduction

Over recent times, output[1]-driven funding systems have become a popular funding model for the provision of tax-financed public services in many part of the world. The principle of output-driven funding is not merely a budgeting model for application in the funding by government of public sector service-delivery agencies. It is also one model upon which to base contracting with private sector suppliers for the outsourced production of tax-financed goods and services. Indeed, the introduction of output-driven funding systems within the public sector often represents an explicit attempt to facilitate competition between in-house and outsourced supply, as one means by which to harness powerful market forces in the cause of enhanced productive efficiency.

Output-driven funding systems have been introduced in particular sectors, an important example being the introduction of hospital funding systems in which funding is based in significant measure upon outputs as measured using the so-called 'casemix' system (in the US, 'prospective payments') based upon 'diagnostic related groups' (DRG) output groups. In Australia and New Zealand, there has in recent years been a radical attempt to apply the principle of output-based funding to the entire annual government budget through a system of so-called 'accrual output budgeting' (AOB)[2]. The essence of AOB is the attempt to transform public budgeting into a process whereby government purchases outputs from its agencies, under circumstances intended to emulate the competitive market. The aim is to make

* School of Economics and Finance, Queensland University of Technology, PO Box 2434, Brisbane, Qld, 4001, Australia, Phone: + 61 7 33009770, Fax: + 61 7 35110459, Email: m.robinson@qut.edu.au

[1] The primary meaning of the term *output* is a good or service delivered to the public. In some jurisdictions, the concept of an output is extended to include one or both of the following: firstly, services (such as policy advice) delivered to ministers, cabinet and the legislature and, secondly, services delivered by government agencies to other agencies rather than to the public (what at the Commonwealth level has been referred to as 'enabling outputs'). *Outcomes*, on the other hand, are the changes induced by the output upon the characteristics of people, social structures, or the physical environment. For example, foods are outputs, and well-nourished people their outcome. Rural advice services are outputs, and higher agricultural productivity and incomes their intended outcomes.

[2] AOB is the Australian term. The system has been commonly referred to in New Zealand as 'contract budgeting'.

agencies operate like businesses, in particularly by focussing their attention on their operating result (profit/loss) (Schick, 1996; Robinson, 2000; DOFA, 1998)

Notwithstanding their popularity, many questions have arisen concerning the practicability and efficacy of output-driven funding systems, and these are the focus of this article. The structure of the article is as follows. Firstly, the concept of a *pure* output-driven funding system is defined, and contrasted with other results-oriented funding and budgeting models. This is followed by an overview of the way in which *real-world* output-driven funding systems tend to differ from the *pure* model, and of certain problems widely identified with the output-driven funding principle. This leads to discussion of the way in which the nature of many public sector outputs affects their suitability of output-driven. The main focus here is upon the implications of output heterogeneity. However, the implications of the public sector focus upon service outputs (as opposed to physical goods) and on the role of 'contingent capacity services' are also discussed. The article concludes with observations concerning the scope for the use of output-driven funding models in the public sector.

The analysis is general in nature. However, there is a particular focus upon AOB by way of concrete illustrations of key points. This reflects not only the authors' recent research focus, but also a view that the ambition of the AOB program for placing the *entire* budget sector upon an output-funded basis makes it particularly relevant to the present topic.

What is an output-driven funding system?

We can define a *pure* output-driven funding system as one that possesses two properties. The first and most essential of these is that the funding provided, or payment made, to a service-delivery agency is an function exclusively of the quantities of outputs (of defined quality) delivered by the agency, with the single permissible exception of funding provided to cover capital costs or capital expenditure. The simplest version of an output-driven funding system is one in which a specific, constant per-unit 'price' is paid for each type of output delivered by the agency (eg funding a hospital $x for every hip fracture treated, and $y for every ruptured appendix treated), and the agency is expected to cover all of its costs (fixed as well as variable) from this price. Although the discussion that follows will assume the simple constant unit-price version of output funding, this is purely for ease of exposition. The analysis is fully applicable to output-driven funding systems based upon less simple functional relationships between output quantities and funding. For example, an alternative model of output-driven funding is one in which lump-sum funding of capital costs is combined with a constant per-unit price from which the agency would be expected to at least cover its variable (but not its fixed) costs[3].

The second defining characteristic of a pure output-driven funding system is that in such a system the aim is to base the output 'price' upon what might be called the efficient cost of production[4]. Thus in a casemix hospital funding system the prices paid for the treatment of the relevant outputs typically relate to the cost levels achieved by the more efficient hospitals. Accrual output budgeting systems have proposed that this principle be generalised as the

[3] Similarly, it is perfectly compatible with the principle of output-driven funding for the per-unit 'price' (ie average revenue) to fall as quantity increases, perhaps as a reflection of marginal costs below average variable cost.

[4] This concept is logically most suited to the determination of a 'price' that is intended to cover fixed as well as variable cost, by reference to the efficient total cost of production. However, it has also in practice been applied to the determination of 'prices' intended to cover variable costs.

basis for funding all outputs produced by budget sector agencies (Department of Finance and Administration, 1999: 27; Victorian Department of Treasury and Finance, 1999: 9, 11; Queensland Treasury, 1998: 23).

An implication of the first property is that, in a *pure* output-driven funding system, the principle of 'payment for results' applies. This means that, just as in a simple market transaction such as the ordering of office supplies, payment is made only for products actually delivered. If the supplier delivers less than has been ordered, the payment made will be reduced accordingly.

Applied in a budgetary (as opposed to outsourcing) context to budget-dependent government agencies, a pure output-driven funding system has the quite deliberate effect of introducing the business distinction between the costs of production and the revenue earned from the sale of products (eg Ball, 1992: 21). With output 'prices' based upon efficient cost, an agency which produces inefficiently so that its average cost exceeds the price it receives would, just like an inefficient business, make a financial loss. This would encourage the agency either to lift its game or to outsource production to an efficient private sector producer. Similarly, if the principle of 'payment for results' were to apply, agencies would not be able to respond to budget cuts by reducing output quantity or quality, because this would simply result in a commensurate further reduction in funding. The result would be (in the language of New Zealand contract budgeting) the transformation of 'budget pressure' into 'purchase pressure'.

It is the cost/revenue distinction that explains the introduction within the accrual output budgeting framework of full business-style operating statements with a bottom-line profit/loss measure (the 'operating result'). The notion that the operating result can be a meaningful overall performance indicator for a budget-dependent government agency is fundamentally at odds with the traditional view of government budgeting. In the traditional view, budget funding is not considered to be revenue earned by the agency, but rather to be more in the nature of a spending allowance. Any reference to *agency* revenue (as opposed to government revenue) is, from the traditional point of view, understood to refer only to any 'user-charge' revenue raised directly from the public by the agency itself. This corollary of this is that budget-dependent agencies are considered to be as inherently loss-making, and there is thus no question of the operating result serving as a business-style performance measure.

This makes it clear that the AOB objective of making a business-style profit/loss result a key focus of agency management is not simply a by-product of the adoption of accrual accounting. Accrual accounting can be introduced into the budget sector of government without the use of profit measures, as has occurred in the Australian State of New South Wales[5]. There are management accounting and fiscal policy grounds for adopting accrual accounting in the budget sector which have nothing to do with agency-level profit measures (Robinson, 2001, 2002a).

The 'market' inspiration of output-driven funding systems is obvious. However, real-world market structures show great diversity, and certainly do not all confirm to the model implicit in output-driven funding systems (Williamson, 1985). The specific type of market model

[5] NSW uses for budgetary purposes truncated 'operating statements' which do not report an operating result, but rather the 'net cost of services' (ie expenses minus agency own-source revenue). It is true that operating statements in departmental annual reports do report operating results, but this is driven merely by the desire to comply fully with accounting standards.

upon which these systems are based is the perfect competition model of elementary economics textbooks. It is from this model, for example, that the idea of 'price' based upon efficient cost derives[6]. Other assumptions of the perfect competition model—in particularly, that products are homogenous (standardised)—are also important to understanding the output-driven funding model, as discussed below.

Output-driven funding systems versus other results-oriented funding models

The concept of a pure output-driven funding system is clarified further by noting how it relates to results-oriented funding models generally, including *performance budgeting*.

Results-oriented funding models may be defined as those in which there is a significant linkage between funding provided and the results expected of the funded agency. Although output-driven funding systems fit this definition, so do other models which take a difference approach in respect to either or both (i) the type of 'result' upon which funding is based, and (ii) the nature of the linkage between results and funding.

In relation to the first point, there are funding systems that differ from output-driven systems in only one crucial respect: that they pay a price for outcomes as well as (or instead of) for outputs. Thus, for example, in the Australian Government's Jobs Network scheme for the outsourced delivery of labour market programs, a significant element of remuneration to service delivery agencies takes the form of success fees pursuant to placement of long-term unemployed clients in jobs.

The term *performance budgeting* refers to results-oriented funding models that operate in a budgetary context rather than an outsourcing context. The core idea of performance budgeting is that there should be an explicit linkage between the funding provided to a public sector agency by government and the results that the agency is expected to deliver to the community. This idea is contrasted with 'traditional' government budgeting, where budget funding was allocated on the basis of inputs (eg staff, equipment etc) requirements.

Accrual output budgeting constitutes one specific version of performance budgeting. However, there are many other versions of performance budgeting in force around the world. Both AOB and these other versions of performance budgeting can be said to take their starting point from program budgeting. The core principle of program budgeting is that funding should be allocated to output classes, so as to make the budget an instrument of rational expenditure priority setting. Most contemporary forms of performance budgeting seek to go beyond program budgeting in tightening up the linkage between results and funding, and in nearly all cases this involves the stipulation of explicit output/outcome targets linked to budget funding. However, only in the case of AOB is the aim to develop an output-driven funding systems in the specific sense that as we have defined it in this paper. The other versions of performance budgeting do not pretend to set budget allocations by multiplying output 'prices' by quantities, nor do they generally regard agency profit results as a key performance measures. To take one US example, in the State of Florida under a system of 'performance-based program budgeting' the annual budget acts passed by the legislature include for each agency a set of output and outcome targets (Florida Office of Program Policy Analysis and Government Accountability, 2000). In the United Kingdom, there has

[6] As anyone who has undertaken an introductory economics unit knows, under conditions of long-run equilibrium in perfectly competitive markets, competition between firms and free entry/exit drive price down to minimum long-run average cost.

developed since 1998 a system of Public Service Agreements and Service Delivery Agreements between the Government and agencies linked to the budget (HM Government, 2000). In Australia, the only state not to have adopted accrual output budgeting—New South Wales—has over recent years been progressively introducing a system of Service Resource Allocation Agreements between agencies and executive government (NSW Treasury, 2000).

Output-driven funding systems in the real world

The idea of a 'pure' output-driven funding system is an analytically-useful archetype to which few real world funding system or contractual outsourcing arrangements fully correspond. The point is not merely that many outsourcing contracts include elements of outcome-based funding. In fact, once one steps beyond the realm of purchasing arrangements for simple products like stationary supplies, even supposedly output-driven systems typically include significant elements of *input*-based funding or of funding related to workload levels of support activities. For example, casemix hospital funding systems commonly fund intensive care units on more or less an inputs basis, and supplement output-based 'case payments' with additional per-diem funding for long-stay 'outlier' cases.

Similarly, in the case of accrual output budgeting systems, there tends to be a significant gulf between the theory and the principle of output-based funding. Discussions with key public sector personnel at the State level in Australia suggests that in most States budget allocations are almost never decided *even in part* by multiplying planned output quantity by output prices, notwithstanding all the official statements of intent in this respect. Instead, the time-honoured principles of incremental budgeting—in which last year's budget 'base' is adjusted in line with changes in costs and responsibilities and rough-and-ready treasury perceptions about the scope for efficiencies—continue to operate.

The national government in Australia has made a much more serious and noteworthy attempt to apply the principle of output-based funding through so-called 'purchase agreements' which are negotiated between each agency and the Department of Finance and Administration (DOFA). Under these agreements, a significant portion of the funding provided to agencies comes through per-unit 'prices' paid for a range of output-types[7]. Alongside this output-price funding, there is also a large component of 'base' funding provided, which is not an explicit function of output quantities.

This Commonwealth system is one from which many jurisdictions around the world could learn much. However, the key point here is that this system represents, not a pure output-driven funding system, but rather a compromise between such a system and traditional budgeting. This is because the 'base' funding provided in the 'purchase agreements' does not by any means represent only funding for fixed costs. Many of the outputs produced by Commonwealth agencies are not funded, even in part, through per-unit 'prices', and funding for both variable and fixed costs associated with these outputs is in effect included in the base funding components of agency funding.

In those cases where outputs are funded partly via a per-unit price, it would seem that notion of setting price on the basis of *efficient* cost has had limited application in practice, even at the Commonwealth level. A practice has developed so-called 'pricing reviews' of

[7] For example, in the case of the Department of Immigration and Multicultural Affairs, specific per-unit prices are paid for each of approximately twenty categories of visa applications outputs (and prices are specified also for a number of other departmental outputs).

departmental budget allocations are conducted jointly by DOFA and targeted departments[8]. However, these appear to have made only selective use of output cost benchmarking process and related techniques for measuring 'efficient costs'. Indeed, some might question the extent to which these 'pricing reviews' differ from pre-AOB agency-level budget reviews.

To complete the story about AOB, it should also be mentioned that there have been only limited attempts to give effect to the principle of 'payment for results'. In state jurisdictions employing AOB, agency budget allocations are explicitly based upon *planned* output quantities, without any intention of adjusting funding when and if the output quantity actually delivered differs from that planned. The only exception to this has been in Victoria, where an ultimately unsuccessful attempt was made to put into practice the principle of payment for results (Robinson, 2002b).

Turning from the specifics of AOB to output-driven funding systems generally, it is clear that when the basic principles of output-driven funding are put into practice, significant issues tend to arise. In particular, there are recurring concerns that the focus upon outputs in such systems can result in reduced attention to outcomes and to a focus upon output quantity rather than quality. In particular, many worry that attempts to apply 'purchase pressure' by cutting the output 'price' can all too easily, and contrary to the intention of the model, erode the effectiveness of the services being provided. A particularly graphic illustration of these concerns arose when the casemix hospital funding model was hastily implemented by the incoming Kennett Government in the Australian State of Victoria, in the context of major cuts to hospital budgets. The result, it is now generally accepted, was a quite serious erosion of the quality of service provided in Victorian hospitals (Productivity Commission, 1997; Victorian Auditor-General, 1998).

These comments are not meant to belittle the real achievements of AOB or of other output-driven funding systems. However, they do raise the question of whether there are fundamental limits and constraints on the extent to which one can efficaciously apply in the public sector the principle that funding should be based on output quantity.

In the Australian AOB context, there is a view in that the difficulties which have been experienced in implementing the output-driven funding principle are essentially transitional and technical in nature. Thus, for example, the 1999 *Review of Budget Estimates Production Arrangements* commissioned by DOFA and the national Treasury suggested that once 'robust management accounting systems' were developed and outputs/outcomes better specified, AOB would be able to operate as the output-driven funding system it was intended to be (Vertigan, 1999: 3). It is argued below that this seriously understates the nature of the problems which arise in implementing output-driven funding models. It is suggested that, when the nature of budget-sector outputs is properly considered, it becomes clear that, although the principle of output-driven funding suits some public-sector outputs, there are many other outputs for which it is either inherently inappropriate or problematic. Moreover, where the principle does or can work it is almost never along the simple lines of the 'pure' model outlined above. The compatibility of quality services with output-driven funding systems in considered in this context.

[8] Outlines of the results of many of these can be obtained from the Portfolio Budget Statements of affected agencies, which many be accessed from the annual budget documentation at the Commonwealth budget Website (http://www.budget.gov.au/pbs/default.htm).

Why per-unit output funding often doesn't work

Because an output-driven funding system is one in which a certain per-unit 'price' is paid for each type of output[9], such a system requires a clear definition of what constitutes an output-type for funding purposes and also of what constitutes a unit of each of those types of output. In the simple textbook version of the market which provides the template for output-funding models, there is no ambiguity about product-types and units of product. This is because the textbook world is one of mass-produced homogenous (standardised) physical products. A nail manufacturer will, for example, produce standardised types of nails: for example, 5 cm nails, 3 cm nails and 1 cm nails. Each type of nail will have its own price. What makes the 5cm nails a distinct product type for marketing and pricing purposes is the fact that each 5 cm nails is essentially identical to other 5 cm nails (and different from all 3 cm and 1 cm nails). Equally, defining the *unit* of output is straightforward because it is easy to define the characteristics which distinguish a completed nail from the semi-finished product.

Unfortunately, defining output-types and units of output is often nowhere near as simple as this in government. Budget-sector outputs are mainly services rather than goods, and they are generally not standardised. Consider, for example, a program (such as the current Australian SAAP program) that provides emergency assistance to women who are victims of domestic violence. Suppose that, at least as a first iteration, this service is considered to be a single output-type for funding purposes so that, under an output-funding model, service-delivery agencies would receive a specific price ($z) for every domestic violence assistance case handled. The obvious problem with this is that there are considerable variations in the extent of service required per case, because of differences in the circumstances and needs of clients. At one end of the spectrum, some clients may need nothing more than a couple of night's emergency accommodation. At the other end of the spectrum, those with the fewest personal resources (financial, social and emotional) may need extensive counselling, advocacy services, and accommodation for an extended period. As a consequence, there will be considerable variations in the cost-per-case, for reasons that have nothing to do with efficiency. There may be many cases which will cost the agency more than the fixed price of $z per case, and equally other cases which cost less than $z.

Output heterogeneity is the general term used to describe such differences in the level of services, and therefore in costs per unit of output, which arise from differences in client or case characteristics between units of the 'same' type of output. Output heterogeneity is very widespread in government. Significant output heterogeneity creates major problems for output-driven funding systems. It is, as discussed below, a key reason for the quality maintenance problems which can arise under such systems. Setting the quality issue aside for the moment, however, an obvious problem concerns price setting: how does one set the output price for an output the costs of which vary greatly between one output unit and another? The least unfavourable circumstances one may face in this respect are those where, although an output is characterised by considerable heterogeneity, the *average* cost per unit of output is stable. If this happens to be the case, setting an output price based on average cost may make an output-driven funding system workable *if* the service-delivery agencies is prepared to adopt a 'swings and roundabouts' approach to output heterogeneity. This means that the agency accepts that, although it will lose money each time it delivers a high-cost unit of output, these losses will be offset by profits made delivering lower-cost units of output.

[9] As mentioned in note 2, this is a simplification, and the analysis holds more generally.

This points to the other key problem to which significant output heterogeneity gives rise in an output-driven funding system. There is a clear incentive in such a system for the service-delivery agency to avoid high-cost cases and to seek out lower-cost, profitable clients and cases. The usual term for this practice is *cream-skimming*. It is a practice which is incompatible with widely-held principles of equity in respect to access to key public services. Cream-skimming tends to be thought of as a consequence of hard-nosed profit-maximisation strategies employed by private enterprise, and it is probably true that the problem of cream-skimming will generally be greater when service delivery is outsourced to for-profit entities. However, even if public sector or non-profit service-delivery agencies are less driven to make profits, they do have a strong incentive to avoid losses. This means that, other than in the special circumstances where the stability of average costs makes the 'swings and roundabouts' approach feasible, such agencies may have no choice but to engage in cream-skimming (or quality erosion, on which see below) under an output-driven funding system.

Clearly, outputs characterised by severe heterogeneity are unsuited to output-driven funding. The classic example of the policy advice 'output' illustrates the point. To fund ministerial policy advice by paying an agency $x per 'unit' of advice is clearly impracticable. Policy advice varies in form from oral advice to full formal written 'briefs', with associated cost variability. Even setting different prices for each form of policy advice would not work: there can be huge variations in the amount of effort required to prepare formal policy briefs, depending upon the complexity and novelty of the issues addressed. It is for this reason that policy advice is overwhelmingly financed on an input basis (ie funding a specific complement of policy staff).

There are strategies that can be used to reduce or mitigate cream-skimming and other difficulties for outputs in respect to which heterogeneity is not severe. Government can, for example, reduce or eliminate the capacity of the service-delivery agency to pick and choose amongst client/cases, either by enforcing a 'universal service obligation' or by itself undertaking the client/case selection and referral process. Moreover, the severity of the output heterogeneity problem may be reduced by the use of a larger number of more narrowly-defined output types. Precisely these strategies are employed in the casemix hospital funding model, which arguably shows the way in which mild or a moderate heterogeneity can in the case of some public services be overcome so as to permit the successful implementation of an output-driven funding system (albeit a highly impure one).

Useful as they sometimes are, however, there are great limits on the ability of such strategies to render heterogeneity manageable within the context of an output-driven funding system. It is impossible for reasons of space to give proper consideration here to the nature of these limits, other than to observe that these strategies must frequently confront major informational problems and, consequently, involve considerable (and in many instances, unacceptable) administrative costs. The implication of this is that, even for many government outputs that are characterised by mild or moderate heterogeneity, output-driven funding cannot be expected to pass a cost-benefit test[10].

[10] One key consideration here is the volumes (quantities) of output produced. The larger the quantity of a particular type of output produced, the more likely that the significant information costs of making a output-driven funding system work (such as the costs of accurate output costing, and of benchmarking and other processes used to estimate efficient cost) will be worth incurring. Not only that, but the larger the volume of production, the more stable average cost will tend to be (via what statisticians call the 'law of large numbers'). For converse reasons, subdividing output groups in order to reduce heterogeneity is a process will inherent limitations.

The implications of output heterogeneity can be seen clearly in real-world output-driven funding systems. It is primarily in order to deal with heterogeneity that that the casemix funding system differs from a *pure* output-driven funding system in so many way (eg the arrangements for higher-cost 'outlier' cases noted above). Similarly, it is in large measure because of heterogeneity that under the 'purchasing agreements' introduced under the Australian government's form of accrual output budgeting, there are many agency outputs which are financed, not by per-unit 'prices', but through 'base' funding[11].

Contingent capacity services

Although heterogeneity is probably the most widespread obstacle to the notion of paying a fixed per-unit price for outputs, it is by no means the only obstacle encountered by agencies. What might be called 'contingent capacity outputs' provide another example of a type of service for which the notion of output-driven funding is unworkable. The armed services provide a classic example of a contingent capacity output. The fighting of wars is the principal 'output' of the armed services. Yet the notion that the annual military budget should be determined by paying a certain 'price' per war is patently impracticable. The problem with paying a price per war is not merely output heterogeneity. It is also that, outside an actual war situation, military budgets do not pay for wars. Rather, they aim to maintain the capacity to fight wars should the need arise. Indeed, by maintaining effective armed forces most countries hope to reduce, by deterring potential aggressors, the likelihood that the military will be called upon to deliver the war 'output'.

It might perhaps be thought that wars are simply an unusual and highly atypical type of government output. This is not, however, true. There are quite a few other types of contingent capacity service provided by government. Emergency services such as the fire and ambulance services fall into this category, as do hospital emergency departments (and to some extent also intensive care units). Within agriculture departments, to give just one another example, specialist services are maintained to provide a capacity to respond to outbreaks of serious plant or animal disease. Contingent capacity outputs have three characteristics. Firstly, there is an *essentiality* and *immediacy* to supply—in other words, it is considered important that when a demand for the service arises, the service be supplied promptly. Secondly, the demand for the service is unpredictably variable. Thirdly, the time lags involved in adding to production capacity (arising from factors such as the specialised nature of the human/physical capital used) exceed the desired service response time.

It is not by accident that many of these types of services are concentrated within the public sector. What contingent capacity services provide is a type of insurance policy that would not be effectively provided in the absence of government intervention. In the language of economics, contingent capacity services arise in response to a well-recognised form of 'market failure' (incomplete insurance markets).

Outcomes and output quality

Why is it that problems concerning outcomes and output quality may arise in output-driven funding systems? To answer this question, it is crucial to be clear about the relationship between outputs and outcomes, and about the distinction between outcomes and output quality.

[11] Note, however, that even amongst the outputs covered by per-unit prices, there are many which are affected by quite significant heterogeneity, leading to significant problems

Outcomes are the intended impacts of outputs. Output quality, by contrast, refers to the extent to which the characteristics of the product are such as to facilitate the achievement of the intended outcome. The two are not at all the same. Two units of an output that are of equal quality may well produce different outcomes. For example, two students may sit in the same classroom and receive the same level of attention from the teacher yet achieve significantly different learning outcomes. Indeed, some units of output may totally fail to achieve any outcome, for reasons that have nothing to do with quality. Some road trauma victims treated in hospital emergency units will, for example, die even with the best of care. Expressed differently, unsuccessful or partially successful services still constitute units of output[12].

This means that, in 'purchasing' outputs, one cannot also be *directly* purchasing outcomes. Nor, generally, do output-driven funding systems pretend to be purchasing defined outcomes. For that is not the 'market' model. In simple markets, consumers purchase outputs (products) of defined *quality*, not defined outcomes. In a textbook world of standardised physical goods and perfect consumer information, a customer entering a purchase contract with a supplier will have no trouble specifying the physical attributes of the output to be purchased. (This will commonly involve nothing more than specifying the type and model of an off-the-shelf product.) Nor does the consumer who inhabits the textbook world have any difficulty ascertaining upon delivery that the product is of the required quality. In short, the simple textbook market world is one consistent with what has sometimes been referred to as a 'compliance with specifications' concept of output quality.

The problem is that this 'market' model is quite inappropriate in the real public sector world of highly imperfect information and heterogeneous outputs. Most budget-sector outputs are services rather than goods, and quality measurement is certainly more difficult for services than for goods even when they are standardised. However, the more serious problem is that heterogeneity often fatally undermines the notion of pre-specifying output quality attributes. For a standardised product such as a nail, quality assurance simply involves establishing that each nail complies with a standard set of pre-defined specification (relating to length, shape, hardness etc). For heterogeneous outputs, however, an additional dimension of quality comes into play: namely, are the characteristics of the service appropriately tailored to the needs of the client or the nature of the circumstances involved? For example, did a social worker in handling a domestic violence case correctly judge what the needs of the client were (counselling? advocacy? emergency housing?) and offer the most relevant forms of assistance? Did a physician apply the most appropriate medical procedures in treating a patient admitted with severe road trauma and multiple injuries? In short, quality requires appropriate *variations* in output specifications. For many outputs, this means that it will not be possible or practicable to fully pre-specify desired output characteristics. In many cases, much has to be left to the professional judgement of the service-delivery personnel, and there may be ambiguity even after the event as to the quality of the service provided.

The fact that it is frequently impracticable to contractually pre-specify required output quality has far-reaching implications for the public sector. It makes it much harder to uncouple output 'price' from output quality. It is much more difficult to prevent a service-delivery agency which delivers complex, heterogeneous services from responding to a cut in the

[12] This is an important matter to emphasise, because the mistake is not infrequently made of identifying output with *successful* units of service. An example of this can be found in a recent UK Government performance measurement publication which gave "effective medical treatments" as an example of an output (HM Treasury et al, 2001: 8).

output price it receives by reducing quality than it is to catch the supplier of a standardised physical product who delivers a good which does not meet specified quality standards. This in turn implies that it is will often be impossible to emulate the fierce competitive processes of a simple market, in which competition between suppliers of a standardised product drives the price down without threatening quality.

For these reasons, moreover, quality assurance in the public sector can never be simply a 'market' process. Rather, it must inevitably rely considerably upon non-price mechanisms such as:

- Trust and reputation in the context of long-term relationships, whether in-house between public sector employees and government (through, for example, professional and service ethics), or in the context of mutually-advantageous partnerships between external suppliers and government,
- Monitoring and regulation of processes and inputs: for example, through accreditation systems, staff/client ratio requirements and staff training requirements.

It is for this reason that quality assurance processes separate from the purchase process play a crucial part of real-world output-driven funding systems such as casemix hospital funding and AOB.

Digressing a little, the problems of output quality assurance raise one obvious question for purchaser-provider models: why not purchase outcomes directly? Why not, for example, fund hospitals for the health outcomes achieved rather than for patients treated? Why not fund schools based on the educational accomplishments of their students (the 'proximate' outcome), rather than for teaching (the output) or teachers (the input)?

Elements of outcome-related payment (and, less ambitiously, output quality bonuses) certainly do feature in many funding arrangements, and are often useful performance drivers. However, this does not mean that funding based wholly or largely about outcomes is a viable model, notwithstanding the enthusiasm of management gurus like Osborne and Gaebler (1993: 139). In general, outcome-driven funding is even more problematic than output-driven funding. The problem is not merely that outcomes are generally much more difficult to measure than outputs, although that is certainly a crucial consideration. It is also that the variability and uncertainty which affect cost-per-outcome is in general of a much higher order of magnitude than the variability in cost-per-output arising from output heterogeneity and other relevant factors. Take, for example, the police 'output' murder investigations. Output heterogeneity, arising from considerable variations in case complexity, means that the cost per investigation varies enormously. In some cases the murderer is caught red-handed, with minimal police effort, whereas in some other cases enormous investigative effort is invested before the case is either solved or effectively closed. Funding on a price-per-unit-of-output would therefore be entirely inappropriate. Clearly, however, funding based on a price-per-murder-solved basis (i.e. on an outcomes basis) would be *vastly* more problematic. Indeed, one of the main reasons why some murder investigations are effectively closed without being solved is that in those cases the cost of achieving an outcome has turned out to be too high.

Conclusions

The analysis is this paper suggests that the most basic principle of output-driven funding systems—that government funding/payments to service-delivery agencies be determined as

an explicit function of the quantity of outputs of defined quality produced by that agency—is one which is suitable only for a subset of tax-financed public services. As a rule, it is most likely to work for relatively standardised services which are produced in large volumes. Conversely, there are many tax-financed outputs for which output-driven funding systems are either inherently unworkable, or too administratively costly to be worthwhile. These include more heterogeneous services, and contingent capacity services. A key challenge for those who are keen to see a more results-oriented public sector is to identify where and when principles of output (and even outcome) funding can effectively be applied, rather than pursue simplistic agendas which assume that the model is of universal applicability.

This implies that, notwithstanding the considerable appeal of the notion that funding should be driven solely be results, in practice the financing of many public services must inevitably be based at least in part upon the funding of inputs processes/support activities. This analysis also challenges that proposition that it is possible to transform budget sector agencies into business-type entities for whom the profit/loss result is a meaningful overall performance measure in the way it is for private enterprise.

References

Ball, I. 1992. 'The New Approach: Financial Management Reform', *Accountants' Journal*, June, 17-21.

Department of Finance and Administration, 1998. *Specifying Outcomes and Output: Implementing the Commonwealth's Accrual-Based Outcomes and Outputs Framework.* Canberra: The Department.

Department of Finance and Administration, 1999. *Commonwealth Accrual Budgeting Guidelines*, Canberra: The Department.

Florida Office of Program Policy Analysis and Government Accountability, 2000. *PB^2 Status Report*, Report 00-15, November.

HM Government, 2000. *2000 Spending Review: Public Service Agreements*, Command Paper 4808, presented by the Chief Secretary to the Treasury.

HM Treasury, Cabinet Office, National Audit Office, Audit Commission and Office for National Statistics, 2001, *Choosing the Right Fabric: A Framework for Performance Information*.

New South Wales Treasury, 2000. *Financial Management Framework: the Financial Management Framework for the General Government Sector*, Sydney: The Treasury.

Osborne, D and T Gaebler, 1993. *Reinventing Government: how the entrepreneurial spirit is transforming the public sector*, New York: Plume.

Productivity Commission, 1997. 'Casemix Funding of Public Hospitals in Victoria', in *Reforms in Government Service Provision 1997*, Melbourne: The Commission.

Queensland Treasury, 1998. *Managing for Outcomes: Policy Guidelines*, Brisbane: The Treasury.

Robinson, M, 2001. “Accrual Accounting and the Public Sector”, *Economic Papers*, June 2001.

Robinson, M, 2002a. "Accrual Accounting and Australian Fiscal Policy", forthcoming, *Fiscal Studies*.

Robinson, M, 2002b. “Financial Control in Australian Government Budgeting”, forthcoming, *Public Budgeting and Finance*, 22(1).

Robinson, Marc, 2000. “Contract Budgeting”, *Public Administration*, 78 (1).

Schick, Allan, 1996. *The Spirit of Reform: Managing the New Zealand State Sector in A Time of Change*, A Report prepared for the State Services Commission and the Treasury, Wellington: The Treasury.

Vertigan, M, 1999. *Review of Budget Estimates Production Arrangements*, mimeo.

Victorian Auditor-General, 1998. *Acute Health Services under Casemix: a Case of Mixed Priorities*, Melbourne: The Audit Office.

Victorian Department of Treasury and Finance, 1997. *Output Costing Guide*, Melbourne: The Department.

Williamson, Oliver, 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*, New York: The Free Press.